

Evaluation of a flow cytometry method to determine size and real refractive index distributions in natural marine particle populations

JACOPO AGAGLIATE,¹ RÜDIGER RÖTTGERS,² MICHAEL S. TWARDOWSKI,³ AND DAVID MCKEE^{1,*}

¹ Department of Physics, University of Strathclyde, 107 Rottenrow, Glasgow, G4 0NG, Scotland

² Helmholtz-Zentrum Geesthacht, Center for Materials and Coastal Research, Max-Planck-Straße 1, Geesthacht, Germany

³ Harbor Branch Oceanographic Institute, Florida Atlantic University, 5600 US 1 N, Ft. Pierce, FL 34946, USA.

*Corresponding author: david.mckee@strath.ac.uk

Received XX Month XXXX; revised XX Month, XXXX; accepted XX Month XXXX; posted XX Month XXXX (Doc. ID XXXXX); published XX Month XXXX

A flow cytometric (FC) method was developed to retrieve particle size distributions (PSDs) and real refractive index (n_r) information in natural waters. Geometry and signal response of the sensors within the flow cytometer (CytoSense, CytoBuoy b.v., Netherlands) were characterized to form a scattering inversion model based on Mie theory. The procedure produced a mesh of diameter and n_r isolines where each particle is assigned the diameter and n_r values of the closest node, producing PSDs and particle real refractive index distributions (PRIDs). The method was validated using polystyrene bead standards of known diameter and polydisperse suspensions of oil with known n_r , and subsequently applied to natural samples collected across a broad range of UK shelf seas. FC PSDs were compared with independent PSDs produced from data of two LISST-100X instruments (type B and type C). PSD slopes and features were found to be consistent between the FC and the two LISST-100X instruments, but LISST concentrations were found in disagreement with FC concentrations and with each other. FC n_r values were found to agree with expected refractive index values of typical marine particle components across all samples considered. The determination of particle size and refractive index distributions enabled by the FC method has potential to facilitate identification of the contribution of individual subpopulations to the bulk inherent optical properties (IOPs) and biogeochemical properties of the particle population. © 2017 Optical Society of America

OCIS codes: (010.4450) Oceanic optics; (120.4640) Optical instruments; (290.4020) Mie theory; (290.5850) Scattering, particles; (350.4990) Particles.

<http://dx.doi.org/10.1364/AO.99.099999>

1. INTRODUCTION

Natural particle populations comprise a wide range of material types and sizes. Bulk optical properties like volume light absorption and scattering coefficients are influenced by the summed contribution of all particles in the population with Davies *et al.* [1] recently suggesting the relevant size range extends from 0.01 μm to the mm size class. Whilst instrumentation exists to determine particle size distributions, including the Sequoia LISST series of instruments for in situ measurements, direct observation of material composition is less easily available. This study revives a technique pioneered by Ackleson and Spinrad [2] to determine both particle size and refractive index distributions and assesses the quality of such observations against known material properties for lab samples and independently determined data for natural samples.

Flow cytometry (FC) was originally developed for biomedical studies, and is still used in this fashion in a large majority of its applications. An estimation made a few years ago by Dubelaar & Jonker [3] put the figure for biomedical applications at about 95% of the total, with the rest divided between fields as disparate as pharmaceutical industry, dairy industry, food and water quality control, botany and, of course, marine science; there is little reason to think the figure has changed much in the intervening years. Within marine science, a large number of studies employ flow cytometry in a traditional fashion, using fluorescent dyes and DNA stains to label cells and study their physiology or exploiting in-line imaging to conduct monitoring and taxonomy assessments of marine nanoplankton populations [4-9]. When it comes to marine optics, however, all components of a population of suspended particles contribute in some degree to the optical behaviour of the water, be they living cells, dead cells, organic detritus or indeed inorganic particles and suspended sediments. Therefore, while fluorescent signals are the most commonly used

parameters for triggering particle detection and are certainly well suited when phytoplankton is the focus, they are also insufficient when the entirety of the particle population is considered.

The flow cytometric method for particle diameter and real refractive index determination described here (FC method) follows in the steps of Ackleson & Spinrad [2] and more recently Green *et al.* [10,11] by using the scattering signal (specifically side scattering) as trigger instead, and combines Mie theory and flow cytometry to model the output of the sensors of the flow cytometer and assign size and refractive index to each individual particle. A broadly similar approach has recently been shown to produce useful estimates of scattering cross sections for silica beads and algal cultures [12,13]. This study presents a detailed description of the technique to derive size and refractive index distributions from flow cytometry data and an assessment of performance against known standards and independent measurements taken on natural water samples.

2. SIZE AND REFRACTIVE INDEX DETERMINATION FROM FLOW CYTOMETRY

A. Mie theory and VSF modelling

A detailed description of Mie theory [14] and its derivation can be found in van de Hulst [15] and, more recently, Bohren & Huffman [16]. For the purposes of the FC method, the relevant equations are those for the volume scattering function (VSF) β

$$\beta(\theta, \lambda) = \frac{1}{2} \frac{1}{k^2 r^2} \frac{[i_1(\theta) + i_2(\theta)] I_i(\lambda)}{E_i(\lambda)} = \frac{1}{2} \frac{[i_1(\theta) + i_2(\theta)]}{k^2} \quad (1)$$

and the spectral scattering coefficient b

$$b_{\theta', \theta''}(\lambda) = 2\pi \int_{\theta'}^{\theta''} \beta(\theta, \lambda) \sin(\theta) d\theta. \quad (2)$$

Here θ' and θ'' are angles determined by the aperture of the sensors within the instrument, $k = 2\pi/\lambda$ is the wave number, I_i is the incident intensity, and E_i is the incident irradiance. $i_1(\theta)$ and $i_2(\theta)$ are the scattering amplitude functions; these are core quantities to be determined in any Mie scattering problem, and are fundamentally dependent on the diameter and refractive index (n) of each particle [15].

Mie theory calculations can be carried out for wide ranges of particle diameters and refractive indices and produce a look-up table of simulated forward scattering (FWS) and side scattering (SWS) weighted for sensor-specific shape characteristics. After a correspondence between simulated and measured scattering values is established using standard particles of known diameter and n_r , each new measurement pair of forward and side scattering is mapped against the closest node in the look-up table and assigned a corresponding diameter and n_r . The resulting particle data is then binned to produce particle size distributions (PSDs) and particle real refractive index distributions (PRIDs).

This technique is fundamentally different from the traditional particle sizing method a flow cytometer like the CytoSense is designed to employ. The incident laser beam within the CytoSense is shaped by a cylindrical lens into a thin blade of light, 5 μm thick and 300 μm across. The length of a particle is then assessed from the width of the signal curve produced as the particle transits through the incident laser beam and scatters its light into the sensors (signal baseline length or full-width at half-maximum, FWHM, of the curve). The laser beam is made to be thin to ensure high length and structural resolution; nevertheless,

the traditional technique cannot discriminate sizes smaller than the 5 μm thickness of the laser beam and does not provide n_r information.

The FC method employs an independently developed MATLAB script, using the FASTMie code developed by Slade [17] to handle calculations of the scattering amplitude functions $i_1(\theta)$ and $i_2(\theta)$. The imaginary refractive index (n_i) was kept fixed and set to zero: a fixed value for the imaginary component is necessary as the method does not offer information on this term, and was appropriate for the polystyrene standard beads which were used for tuning the model. Test simulations for realistic n_i values suggested negligible effect of this assumption on diameter and n_r identification. The quantities $i_1(\theta)$ and $i_2(\theta)$ were used to calculate the single-particle VSF of eq. (1) for all combinations of 40 n_r values ranging from 1.335 to 1.725 (relative to vacuum) in increments of 0.01, and 300 log-spaced diameter values ranging from $1\text{e-}8$ to $1\text{e-}4$ metres. The angular resolution for the VSF varied between 0° and 180° , with 0.01° increments between 0° and 1° , 0.1° increments between 1° and 10° and 1° increments between 10° and 180° , for a total of 361 angular values. Finally, $\lambda = 488$ nm (for vacuum) was used for the VSF calculations conforming to the wavelength of the laser source within the CytoSense. This wavelength was further corrected for transmission through water using the absolute (approximate) n_r value for water (1.33). All particle n_r values in the following are given relative to this value unless otherwise stated.

Once obtained, each VSF was integrated to simulate scattering within the flow cytometer, and the results scaled to establish correspondence with measured scattering from particles of known size and refractive index, thus producing the required look-up table.

B. Sensor shape weighting functions

Simulation of the FC scattering sensors requires characterisation of the correct $[\theta', \theta'']$ angles over which eq. (2) is to be integrated. Sensor angle ranges were first manually determined by rough approximation. Improved ranges were then refined iteratively until the model grid of look-up table nodes matched scattering data from polystyrene standard beads of known diameter and refractive index. The VSF given here is axially symmetric along the direction of the incident light, so that integrating over an angular range accounts for the scattering into ring sections of the radiant sphere. The sensors however only intercept part of the ring sections, and weighting functions are therefore needed to further account for sensor shapes and their relative weights.

The forward scattering weighting function (polar angle range: $\pm[2^\circ, 9.7^\circ]$) is defined by the geometric setup shown in Fig. 1a. The particular shape is caused by a vertical "beam stop", in fact a gap between the prisms used to direct FWS light towards the photodiodes. The length of the arc $n = m \cdot l$ for each point in the range $[\theta', \theta'']$ is to be determined and used as a correction factor after normalisation to the corresponding arc x of a unit circle. From trigonometry, it derives that the forward scattering weighting function is expressed as

$$x_{\text{SHAPE},f}(\theta) = \frac{\pi / 2 - \arcsin\left(\frac{\sin \theta'}{\sin \theta}\right)}{\pi / 2}. \quad (3)$$

The shape of the forward scattering weighting function is shown in Fig. 1c. Note that the weighting function is given here for just a quarter of the actual area; correct proportionality is however ensured by scaling factors which will be presented in the following.

The side scattering weighting function (polar angle range: $[45^\circ, 135^\circ]$) is defined by the geometric setup shown in Fig. 1b. The length of the arc x for each point in the range $[\theta', \theta'']$ centred on $\theta = 90^\circ = \pi/2$ is the parameter to be determined. The final form of the side scattering weighting function is expressed as

$$x_{SHAPE,s}(\theta) = \frac{\arcsin \left[a_{\max} \sin \left(\pi \frac{\theta - \theta'}{\theta'' - \theta'} \right) \right]}{\pi / 2}, \quad (4)$$

where a_{\max} is the longest half-chord of the circular projection of collected light, perpendicular to the θ -axis. The shape of the side scattering weighting function is shown in Fig. 1d.

Similarly to forward scattering, the side scattering weighting function is given here for just half of the actual area, but correct proportionality is once again ensured by scaling factors which will be presented in the following. Having obtained the expressions for both sensor weighting functions, for a single wavelength and axially symmetric scattering the final form of the modelled FWS and SWS within the flow cytometer is given by

$$b_{grid}^i = \int \beta_p(\theta) x_{SHAPE,i}(\theta) d\theta, \quad (5)$$

where i denotes either forward or side scattering.

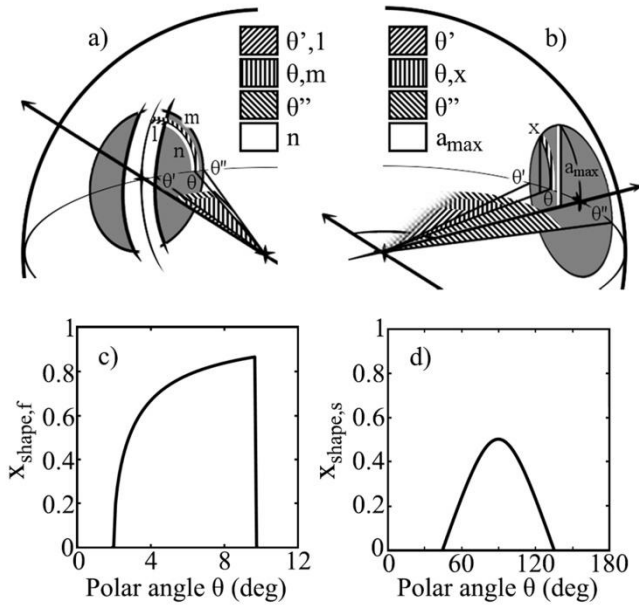


Fig. 1. (a) General geometric setup for the forward scattering shape correction calculations. Angles θ' , θ and θ'' are highlighted in diagonal right, vertical and diagonal left stripes respectively, with corresponding arcs l and m in diagonal right and vertical stripes and arc n in white. (b) General geometric setup for the side scattering shape correction calculations. Angles θ' , θ and θ'' are highlighted in diagonal right, vertical and diagonal left stripes respectively, with corresponding arc x in vertical stripes and a_{\max} in white. (c) Shape of the forward scattering weighting function. (d) Shape of the side scattering weighting function.

C. Grid scaling

Having established relevant angular ranges and respective weighting functions, simulated particle VSFs were integrated to model the output of the flow cytometer, producing a grid of isolines of diameter and real refractive index connecting the look-up table nodes (Fig. 2). The model was centred and scaled on instrument output data for standard particles of known diameter and refractive index. The average forward and side scattering measurement data for 0.5 μm polystyrene beads (Duke Particle Counter Size Standards, Thermo Fisher Scientific for these and all other polystyrene beads unless otherwise noted) was used as reference for the centring of the model

grid. The n_r of the beads is given as 1.197 relative to the approximate n_r value of water (absolute n_r value for polystyrene beads is certified by NIST for $\lambda = 589$ nm, but the effects of wavelength dependence on n_r are small, and were ignored here). After scaling the distance between the corresponding 0.5 μm and 1 μm points in the model grid to be the same as the distance between the measured 0.5 μm and 1 μm polystyrene bead averages, the 0.5 μm point in the scaled model grid was centred on the 0.5 μm polystyrene bead average, i.e.

$$x_{GSCALE}^i = \frac{b_{avg,i}^i - b_{avg,0.5}^i}{b_{grid,i}^i - b_{grid,0.5}^i} \quad (6)$$

$$b_{grid}^i = x_{GSCALE}^i (b_{grid}^i - b_{grid,0.5}^i) + b_{avg,0.5}^i, \quad (7)$$

where i denotes either forward or side scattering, b_{grid}^i the scattering values of the original model grid, b_{avg}^i the average measured scattering values of the polystyrene beads, and b_{grid}^i the rescaled model grid values. The grid scaling performed via eq. (6-7) accounts for both the proportionality required by the sensor shape weighting functions and unknowns such as the incident intensity of light on the particles and the conversion rate of photons to signal in the sensors. It should be noted that in the case of large real refractive indices the n_r isolines produced by the scattering model tend to bundle up together, particularly in the ~ 2 -10 μm range: this will have consequences on the detection of the real refractive index of the particles, as will be shown in the following.

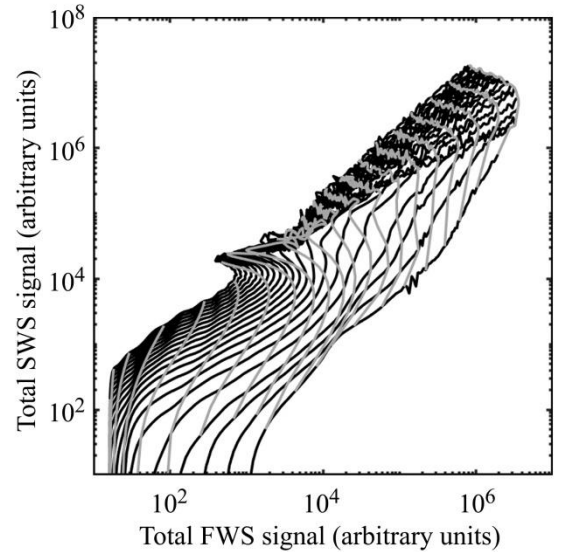


Fig. 2. Scattering model grid produced by the FC method, with isolines of diameter (grey) and isolines of real refractive index (black). For clarity purposes, fewer grid lines are shown than were actually used in calculations.

D. PMT sensitivity

The strength of the SWS signal measured by the flow cytometer is modulated by the sensitivity setting of the photomultiplier tube (PMT) used as side scattering sensor. The range of particle sizes that can be measured with a single PMT sensitivity setting is limited: low sensitivity settings let the instrument detect larger particles without saturating the side scattering PMT, but fail to capture smaller particles; high sensitivity settings let the instrument successfully detect small particles, but hamper the identification of large particles which saturate the side scattering PMT. To overcome this issue multiple PMT

sensitivity settings were used for each sample and the results combined together into a single dataset representative for the whole sample.

Care needs to be taken, as each PMT sensitivity setting needs a separate run (sensitivity run) and corresponding data ranges partly overlap, such that simple summation of the data matrices will artificially inflate measured concentrations. To deal with this, data from each sensitivity run was cut along user-defined thresholds, producing data subsets that were then merged into a single dataset, working under the necessary assumption that the instrument produces mutually consistent runs for each sub-sample. Multiple sensitivity runs with standard beads (1 μm polystyrene beads) were used to scale the sensitivity run data before cutting and merging the total dataset. One sensitivity setting is used as the base setting to which all others are scaled:

$$x_{\text{SCALE},e}^i = \frac{b_{\text{avg},o}^i}{b_{\text{avg},e}^i} \quad (8)$$

where i indicates either forward or side scattering, e the sensitivity setting, $b_{\text{avg},e}^i$ the average scattering value of the standard beads for that sensitivity setting, and $b_{\text{avg},o}^i$ the average scattering value of the standard beads at the base setting. The choice of a PMT base setting in theory doesn't affect the results of this merging; however, a PMT base setting capable of capturing the standard beads necessary for grid scaling without need for interpolation while still keeping PMT saturation in the larger particles at a minimum is preferable in practice. This is suggested as the lowest PMT sensitivity setting capable of clearly determining both 0.5 μm and 1 μm polystyrene beads scattering averages.

Total analysed volumes for each sensitivity run differ even though the total processed volume is the same across all runs. The discrepancy is due to data transfer overhead – the time during which particle data is acquired and sent to the computer and the instrument is not analysing (but still consuming) the sample (CytoBuoy, private communication). This does not produce errors in the estimation of the concentration, but requires further care in the way the final dataset is composed. In this study, the largest total analysed volume among all sensitivity settings was used to scale all other total analysed volumes:

$$x_{\text{VOL},e}^i = \frac{V_o}{V_e} \quad (9)$$

where e once again marks the individual sensitivity settings and V_o is the volume used as reference. This factor was used to adjust the concentrations of the corresponding sensitivity run. Data from all sensitivity runs was subsequently merged into the total dataset, and the total number of particles divided by the largest total analysed volume. The resulting dataset is managed as a single object, but the resulting concentrations are respectful of the individual contribution of each run:

$$\begin{aligned} N_{\text{tot}} &= \sum_e N_e x_{\text{VOL},e} = \sum_e N_e \frac{V_o}{V_e} \\ \Rightarrow C_{\text{tot}} &= \frac{N_{\text{tot}}}{V_o} = \sum_e \frac{N_e}{V_e} = \sum_e C_e. \end{aligned} \quad (10)$$

It should be noted that while in this implementation of the method the largest total analysed volume was used as reference to scale all other total analysed volumes, the choice is indeed arbitrary; any of the total analysed volumes (corresponding to any of the sensitivity settings) may be used.

E. Binning

Once the rescaled and merged dataset is mapped on the centred and scaled model grid binning can take place. Each experimental FWS and SWS data pair is assigned to the diameter and refractive index of the closest node in the model grid. The closest node in the grid is identified as the one that minimizes the Manhattan distance

$$l = |b_{\text{grid},n}^f - b_{\text{exp}}^f| + |b_{\text{grid},n}^s - b_{\text{exp}}^s|, \quad (11)$$

where f and s indicate forward and side scattering respectively, $b_{\text{grid},n}$ is the scattering value for the n -th node in the grid and b_{exp} is the scattering value for the particle as measured by the flow cytometer. Subsequent binning was carried out using 65 logarithmically spaced diameter bins and 40 linearly spaced n_r bins. Bin values for the diameter (in μm) were calculated as $D_i = e^{k_i + 0.16557N}$, with $k_l = -16.87296$, $k_u = -16.70739$ and $k_m = -16.79017$ for lower limit, upper limit and median values respectively and $N \in [0, 64] \subset \mathbb{N}$. The seemingly arbitrary parameters are caused by the adoption (and extension) of the size bins used in LISST-100x instruments. The diameter range so defined covers the 0.05-2000 μm range recommended by Davies *et al.* [1] to account for the entirety of the optically relevant particle population: note that this is in preparation for inherent optical property (IOP) calculations which will be the focus of future work, and that a large number of these bins will be left empty by the method presented here. Bin values for the real refractive index were calculated relative to vacuum as $n_{r,i} = 0.01N + k_i$, with $k_l = 1.590$, $k_u = 1.600$ and $k_m = 1.595$ for lower limit, upper limit and median values respectively and $N \in [0, 39] \subset \mathbb{N}$, before being scaled relative to water. For each sample, the end result of the binning is a $M_{65 \times 40}$ matrix, with each matrix bin containing the particle concentration for a specific combination of refractive index and diameter.

F. Concentration correction

Preliminary measurements carried out on standard polystyrene bead samples hinted that the CytoSense instrument is underestimating the actual particle concentrations when running at low pump flow rate. A further series of measurements was therefore carried out to characterize this underestimation and develop a correction.

The measurements were carried out on two polystyrene bead samples, one a ready-made suspension of 2 μm beads supplied pre-counted by the manufacturer with a nominal concentration of 2000 ± 200 particles per ml (EZY-CAL Microsphere Size Standards, Thermo Fisher Scientific), the other a sample of 10 μm polystyrene beads with unknown exact concentration. The two samples were used to investigate the effect of possible size selection bias in the underestimation of particle concentrations at low pump flow rates. Any correction can indeed be broadly applied to a mixed sample only if no large selection bias between different particle classes is present, with one likely cause for it being the larger inertia and/or tendency for settling of large particles (size selection bias).

The 2 μm polystyrene beads measurement series was carried out for pump flow rates set at 0.99, 1.99, 2.98, 3.97, 4.97, 5.96 and 9 $\mu\text{l/s}$, with ten replicates for all settings but the last. Only seven replicates were made for 9 $\mu\text{l/s}$ due to the large consumption of sample at this flow rate. The results were found to plateau at ~ 3 $\mu\text{l/s}$ without ever reaching the nominal concentration expected for the sample. The expected value of ~ 2000 part/ml was found to be 1.088 times the plateau value obtained as the average concentration of all replicates at 4.97, 5.96 and 9 $\mu\text{l/s}$. This is interpreted as a systematic discrepancy, and the value kept as a correction factor to be applied at all flow rates. Subsequently, all concentrations were normalized using the plateau value (Fig. 3a).

The results indicated a clear relationship between pump flow rate and measured concentration in the sample. Such relationship was found to be well fitted by an exponential curve of the type

$$y = ae^{-bx} + c, \quad (12)$$

suggesting the possibility of using one such function to compensate for the underestimation. The correction curve was obtained via least-squares fit of pump flow rate vs. normalized concentration of the standard beads, providing the means of calculating a concentration correction factor for each particular pump flow rate. In other words, for a specific sample pump flow rate, concentrations are corrected by dividing them by the corresponding resulting factor. Parameters for the 2 μm polystyrene beads correction curve were found as: $a = -0.653$, $b = 1.148$, $c = 0.988$ (Fig. 3a).

The same measurement protocol applied to the EZY-CAL 2 μm polystyrene beads sample was applied to a sample of 10 μm polystyrene beads to investigate the effect of size selection bias on measured particle concentrations, with the same pump flow rate as the 2 μm beads series, starting from 0.99 $\mu\text{l/s}$, and a smaller number of replicates (6 for all flow rates but 9 $\mu\text{l/s}$, for which 4 replicates were made). The actual precise concentration within the sample was unknown a priori, but was not strictly needed as all data were normalized to the value of its plateau (Fig. 3b). The resulting exponential model showed visible differences in comparison with the 2 μm curve. The curve rises more slowly towards the plateau, indicating that larger particles require faster pump flow rates. As a result, the plateau value was obtained as the average concentration of all replicates at 5.96 and 9 $\mu\text{l/s}$ pump flow rates only. Resulting curve parameters were found as $a = -0.634$, $b = 0.496$, $c = 1.020$ (Fig. 3b).

In a compromise between exactness and wider applicability of the method, the differences between the two exponential curve models were considered small enough for the two datasets to be merged and a single correction model to be consequently produced (Fig. 3c), with curve parameters: $a = -0.544$, $b = 0.623$, $c = 1.003$. This is of course an approximation, and in fact density as well as shape of the particles can also be expected to produce a bias in the measured concentrations. Nevertheless, the effect of particle size on measured particle concentrations is the only one that could be assessed reliably, and the loss of exactness caused by the adoption of a generalised model is deemed an acceptable trade-off in exchange for making the method more widely applicable.

This curve was used to correct flow rate-induced concentration underestimations, together with the factor 1.088 for the concentration plateau value described earlier. Their combination produced a correction factor

$$x_{PUMP}(Q) = \frac{1.088}{y(Q)}, \quad (13)$$

where Q is the flow rate and y the value of the general flow rate correction curve. Combining all contributions and correction factors together, following eq. (10) the concentration for each bin in the matrix is given by

$$C_{D,nr}^{tot} = \frac{N_{D,nr}^{tot}}{V_o} x_{PUMP}, \quad (14)$$

where D and n_r are indices of the size and real refractive index corresponding to each bin, respectively. As a final result, summation along the refractive indices or along the diameters provides the PSD and PRID of the sample, respectively.

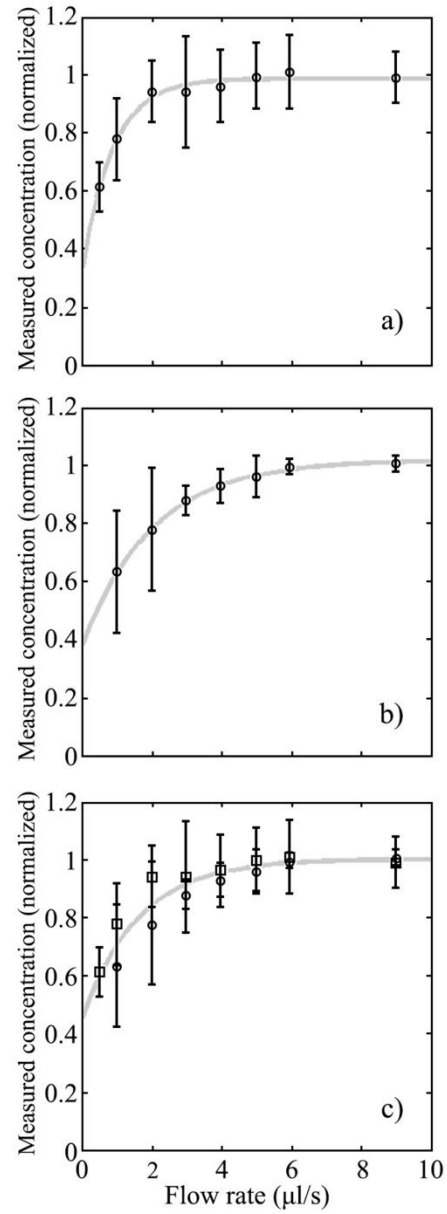


Fig. 3. (a) Concentration correction curve for the 2 μm polystyrene beads normalised concentration values. Curve parameters: $a = -0.653$, $b = 1.148$, $c = 0.988$. (b) Concentration correction curve for the 10 μm polystyrene beads normalised concentration values. Curve parameters: $a = -0.634$, $b = 0.496$, $c = 1.020$. (c) Concentration correction curve for the combined 2 μm (squares) and 10 μm (circles) polystyrene beads normalised concentration values. Curve parameters: $a = -0.544$, $b = 0.623$, $c = 1.003$.

3. RESULTS

A. Diameter and n_r validation

The FC method was first applied to samples containing standard polystyrene beads of known diameter and oil suspensions of known refractive index to validate its correct functioning. To evaluate the overall reliability of diameter retrieval, the FC method was tested on a mixture of polystyrene beads of several different diameters (Tab. 1). The method was found to correctly identify small particle diameters down to the minimum detection limit of 0.5 μm (Fig. 4, Tab. 1).

Polystyrene bead diameter (μm)	
Nominal values	FC values
0.498 ± 0.009	0.496 ± 0.036
0.994 ± 0.015	1.006 ± 0.028
4.993 ± 0.040	5.254 ± 0.433
10.12 ± 0.06	9.266 ± 0.387
50.2 ± 1.0	38.317 ± 1.058
100.0 ± 1.5	63.099 ± 1.243

Table 1. Nominal mean diameter values for the set of polystyrene standard beads used in the validation of the diameter retrieval capabilities of the FC method, and corresponding diameter values retrieved by the method. Error range for both nominal and FC values is given as $\pm\sigma$.

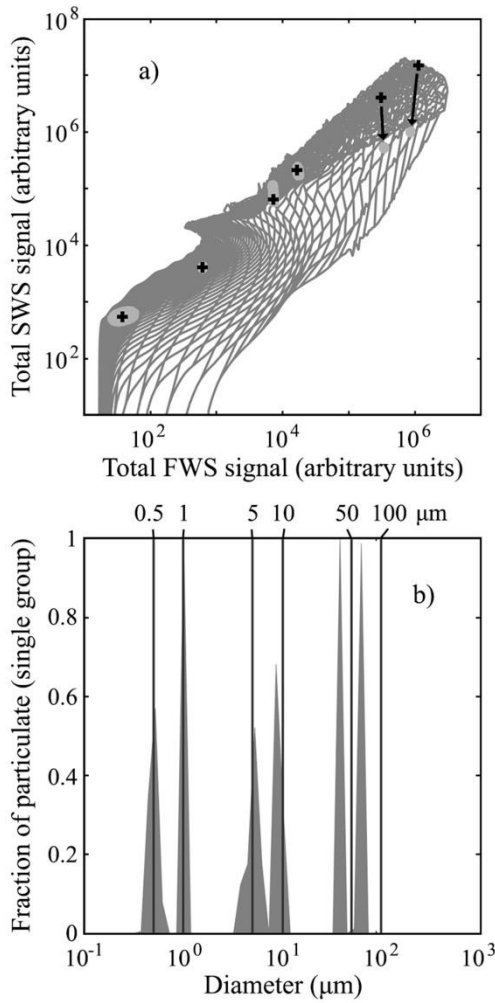


Fig. 4. (a) When data for the set of polystyrene beads of Tab. 1 is superimposed on the grid of Fig. 2 (light grey squares) large particles (upper right area) show a marked discrepancy between measured and expected values. (b) This is reflected in the resulting PSDs, with the FC method increasingly underestimating the diameter of the larger particles ($d > 10 \mu\text{m}$). Target values (cf. Tab. 1) are represented by vertical lines.

Real refractive index			
Nominal values		FC values (relative)	
Substance	Absolute	Relative	Relative
Polystyrene beads	1.595	1.197	1.194 ± 0.002 ($1 \mu\text{m}$)
			1.2 ± 0.037 ($5 \mu\text{m}$)
			1.032 ± 0.002 ($50 \mu\text{m}$)
Olive Oil	1.469	1.102	1.102 ± 0.003
Dodecane	1.421	1.066	1.064 ± 0.002
Nonane	1.405	1.054	1.052 ± 0.002

Table 2. Nominal n_r values for the set of oil suspensions used in the validation of the n_r retrieval capabilities of the FC method, and corresponding n_r values retrieved by the method. Error range for the FC values is given as $\pm\sigma$.

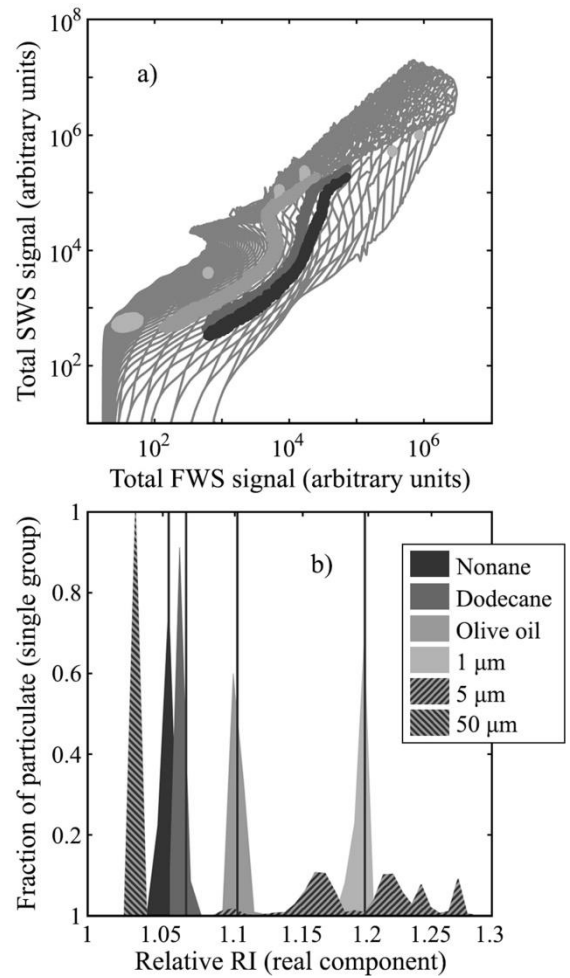


Fig. 5. (a) Data points for the polystyrene beads and three oil suspensions and (b) the resulting n_r distributions. Target values (Tab. 2) are represented by vertical lines. The n_r detection works well at low refractive indices and/or high refractive indices when particles are small, but fails for high n_r values and larger diameters due to the overlap of the isolines of the model grid for these parameters. Real refractive index detection in very large particles produces values which grossly underestimate the expected ones.

Particles larger than 5 μm however presented increasingly anomalous side scattering signal profiles. This is due to the breakdown of Mie theory assumptions i.e. particles becoming larger than the laser beam width (5 μm), a problem that was already recognized by Ackleson & Spinrad in their work [2]. This leads to substantial discrepancies between modelled and measured scattering, causing increasingly marked underestimation of the diameter (Fig. 4b, Tab. 1). Fitting the size peaks with log-normal distributions allowed for the determination of their standard deviation values (σ). The errors on the mean diameter values were subsequently defined as $\pm\sigma$, and found to range between 1.9% and 8.3% of the mean.

It should also be noted that for very small diameters and low refractive indices (bottom left of the scattering model grid) the grid model scaling produces negative scattering values and the model breaks down. This range however lies below the minimum size detection limit of the flow cytometer, and does not affect the results of the FC method at least in its current iteration.

Polydisperse suspensions of oils with known refractive indices (Tab. 2, Fig. 5) were additionally used to test the retrieval of n_r values. The method was found to correctly identify low, organic-like n_r as well as high n_r when particle diameters are smaller than 5 μm (Fig. 5b). However, the tendency for n_r isolines to bundle up together at high n_r values impeded unequivocal determination of the real refractive index of the 5 μm polystyrene beads, and particles with high n_r values around this size range can be expected to provide a generic high- n_r signifier for $n_r \geq 1.15$ as opposed to precise values. Finally, anomalous side scattering profiles for very large particles (e.g. 50 μm polystyrene

beads) was found to lead to gross underestimation of the real refractive index in high- n_r particles (Fig. 5b). As with the size peaks, fitting the real refractive index peaks with normal distributions allowed for the determination of their standard deviation values (σ). The errors on the mean n_r values were subsequently defined as $\pm\sigma$ and found to be $\pm \sim 0.002$ across the n_r range, with the notable exception of the 5 μm polystyrene beads where the bundled n_r isolines impeded the precise determination of an n_r value.

Overall the FC method was seen to be effective in determining the size of particles within the 0.5-10 μm range and n_r values up to ~ 1.15 . Real refractive index values above 1.15 cannot be considered precise, but still positively indicate high refractive indices, while particle diameters above $\sim 10 \mu\text{m}$ are well defined, but increasingly underestimate the actual diameter of the particles. Nevertheless, the 0.5-10 μm particle size range has been recently shown to be responsible for the bulk of scattering and backscattering under typical oceanic conditions [1], and the 1-1.15 n_r range accounts for the most common components of marine particle populations [18,19].

B. UK Coastal Waters (UKCW) dataset

The UKCW dataset consists of natural water samples obtained during the HE442 research cruise in UK waters (4th-21st of April 2015) on board the RV Heincke. Sixty-two stations were sampled across a variety of Case 1 and Case 2 waters around the coast of the UK (Fig. 6), supplying a total of 50 samples with complete sets of FC data, all with matching PSD data from other instruments.

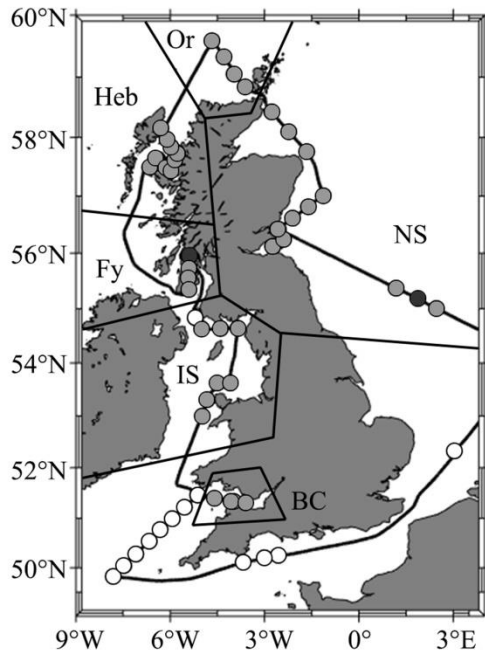


Fig. 6. Track of the HE442 research cruise, which took place in April 2015 in UK waters aboard RV Heincke. Out of the 62 measurement stations visited a total of 50 sets of data were retrieved (light grey circles). Dark grey circles represent stations where two samples were taken. The dataset was divided into area groupings to highlight regional behaviour in the metadata and the measured IOPs. In clockwise order: Bristol Channel (BC), Irish Sea (IS), Loch Fyne and Firth of Clyde (Fy), Hebrides and Skye (Heb), North Atlantic - Orkneys (Or) and North Sea (NS).

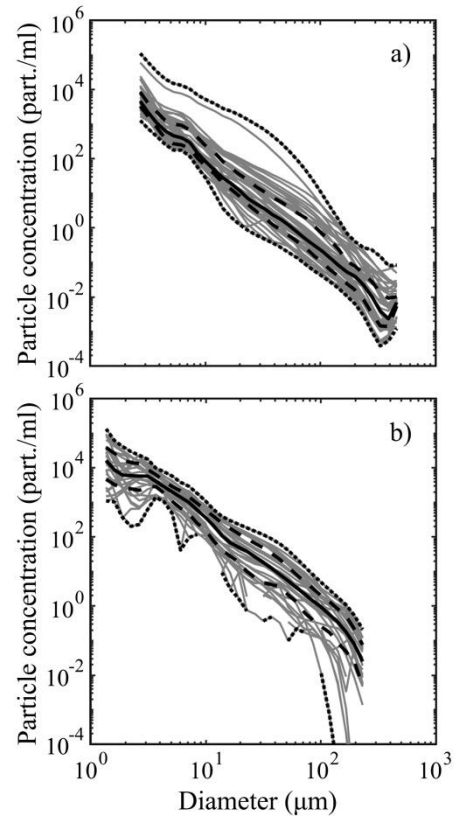


Fig. 7. (a) Collective view of all 50 UKCW PSDs obtained by the LISST-C (grey lines) with overall median, quartiles and max./min. PSDs given as black solid, dashed and dotted lines respectively. In panel (b) the same plot is given for the UKCW PSDs obtained by LISST-B. Note that LISST-B data was available for only 28 out of the 50 samples of the UKCW dataset. Gaps in the minimum distribution are due to gaps in some PSDs retrieved by LISST-B.

This included measurements obtained by two LISST-100x instruments (Type B, size range: 1.25-250 μm and Type C, size range 2.5-500 μm , both by Sequoia Scientific; LISST-B and LISST-C respectively in the following), among other quantities which will not be included in the present analysis. The LISST-100x is a laser diffractometer which allows for rapid acquisition of bulk PSD profiles by measuring scattered intensity in a series of ring-shaped detectors, weighted by the area of each ring [20-22]. Both LISST-C and LISST-B were maintained and operated following the instructions found in the user's manual for LISST-100x instruments [23].

Main depth profiling was done via an instrument frame equipped with Niskin bottles for sample retrieval and one of the two LISST-100x instruments (LISST-C). The frame was lowered into water at each of the stations to measure a profile, and was then kept near surface for water sampling. At a few stations, additional samples from deeper waters were taken. The LISST-C instrument logged data throughout the procedure, and PSD data obtained using the LISST-C was averaged over the stationary phase of the profiling, near the surface or at bottom depth as appropriate to each specific sample (Fig. 7a). Water samples were taken from the Niskin bottles on the frame as quickly as possible after the frame was back on deck and filled into 10-liter plastic containers. In waters with high turbidity the Niskin bottles were

flushed twice to avoid settling out of particulate matter. 48 out of the 50 samples of the UKCW dataset are surface samples (depth: 5-7 m), with two more samples taken from bottom depths instead.

LISST-B was installed on a separate submersible platform, which was deployed and operated independently from the main instrument frame. Depth information from the measurement profiles was compared between the two submersible platforms to extract the LISST-B PSDs matching the data obtained from the other instruments. Two casts were made at each station using the second frame; the LISST-B data used in the following is the result of averaging between the two casts. LISST-B data was available for only 28 out of the 50 samples of the UKCW dataset (Fig. 7b).

All samples were measured by the flow cytometer once for each of four sensitivity settings of the side scattering PMT, for 6 minutes and at a flow rate of 0.5 $\mu\text{L/s}$. Side scattering was used as the trigger channel in all cases. Additional measurements of standard polymer beads for calibration of the FC method were taken daily across the whole sampling period. The low sample flow rate was necessary because of hardware problems which arose in the flow cytometer during the research cruise, causing the sample core to lose stability at higher flow rates.

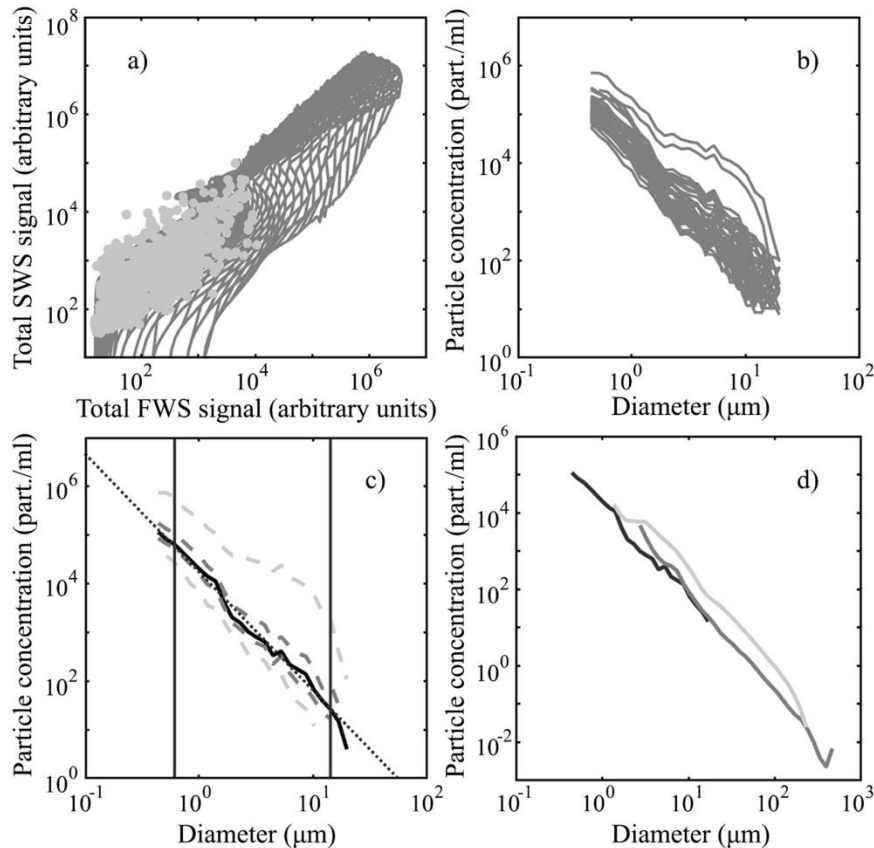


Fig. 8. (a) Example of a typical sample from the UKCW dataset superimposed on the FC model grid and (b) collective view of all 50 UKCW PSDs produced by the FC method. (c) The slope for each PSD was calculated with a power law fit over the available range (minus the first and last two data points to avoid possible boundary effects, as indicated by vertical lines), as demonstrated on the median UKCW PSD (quartiles and max./min. given as dark and light grey dashed lines respectively). In the last panel (d), a comparison between the median FC and LISST PSDs.

C. UKCW results

Fig. 8a shows a typical set of FC data for a natural water sample. In all UKCW samples, the vast majority (>99%) of points were found to lie

within the area of the grid predicted by Mie theory for reasonable estimates of n_r . Points lying outside the grid exhibit properties that diverge from the assumptions underpinning Mie theory, with likely candidates being morphological characteristics such as strong non-sphericity and presence of facets and vertices.

PSDs determined by the FC method were found to broadly follow power law distributions (Fig. 8b). PSD slope values were thus obtained through least squares best fit of power law distributions as defined by

$$N'(D) = \frac{N(D)}{dD} = kD^{-\gamma} \quad (15)$$

after excluding both the first two and last two points in the PSD range to avoid possible boundary effects (Fig. 8c). The form given in eq. (15) is necessary because while the FC PSDs have bin-like nature, PSD slopes in literature are given as those of the underlying density function. Values of the distribution slopes for all 50 samples were found between 2.587-3.813 (mean slope: 3.361 ± 0.250). All PSDs were observed to rapidly lose statistical significance above $\sim 20 \mu\text{m}$ due to few particle counts and particle concentrations reaching a floor corresponding to bins containing a single particle. PSD data above this size threshold was therefore considered unreliable and subsequently ignored. Additionally, PSDs were truncated at their maximum concentration value, generally found at $\sim 0.5 \mu\text{m}$. Concentrations below this size threshold fall spuriously due to the detection limits of the instrument.

Overall consistency in the UKCW PSDs was demonstrated when the dataset was reduced to its overall median, quartiles and max/min PSDs (Fig. 8c), with the exception of two obvious outlier samples that were collected in the turbid, mineral-rich waters of the Bristol Channel (cf. Fig. 8b). These show a clear secondary feature centred on $\sim 8 \mu\text{m}$ that could be associated with strong tidal re-suspension of mineral particles in this area. Fig. 8d, showing median PSDs for the entire UKCW dataset for both FC and LISST instruments, demonstrates broad consistency between FC and LISST results. However, a closer look at ratios of corresponding PSD pairs between the three instruments revealed variability between FC and LISST data as well as between the two sets of LISST data themselves (Fig. 9). LISST-B PSDs were found to be similar to LISST-C PSDs in shape, slope and features as demonstrated by the mostly flat profile of their median ratio (Fig. 9a), as were the FC PSDs and LISST PSDs when compared with each other (Fig. 9b-c). Similarity between FC and LISST PSD slopes was further demonstrated in the shape of the respective distributions of slope values (Fig. 10). As with FC PSD slopes, LISST PSD slopes were determined through least squares best fit of power law distributions after neglecting the extremes of the PSD-LISST overlapping size ranges, which were found to produce boundary effects in the PSD ratios (Fig. 9b-c).

Unlike PSD shape and slopes, particle concentrations were found to differ between the FC and the two LISST instruments, with lower overall concentrations for the flow cytometer and median FC/LISST-C and FC/LISST-B values of $\sim 77\%$ and $\sim 32\%$ respectively across the plateau (Fig. 9b-c). Surprisingly, particle concentrations were also found to differ markedly between the two LISST-100x instruments, with higher overall concentrations for LISST-B and a median LISST-C/LISST-B ratio value of $\sim 48\%$ across the plateau (Fig. 9a).

Real refractive index distributions in the UKCW dataset were found to be fairly homogeneous in shape across all samples (Fig. 11). Highest median concentrations were found in a broad feature between 1.05-1.15, with the overall peak between 1.07-1.09 (peak n_r value: 1.084). This is in accordance with literature values of n_r for the most common components of marine particle populations [18,19], but also smaller than the bulk n_r calculated using the algorithm developed by Twardowski et al. [19] and averaged across all samples (mean bulk n_r : 1.113 ± 0.047). A tail of particles with $n_r > 1.15$ is also observed in all samples, with median concentrations 5 to 10 times lower than peak ones. As discussed previously, real refractive index values above 1.15 are not to be interpreted as precise, but still positively indicate high refractive indices, and the concentration spike visible at ~ 1.3 in

particular is mostly made up of particles above and outside the scattering model grid discussed in. Under the current implementation of the method these grid particles are defaulted at the highest n_r considered, i.e. 1.3, forming the observed feature. These particles were found to account for less than 1% of the total particle population in all samples of the UKCW dataset, and are possibly small, hard inorganic particles with facets and edges for which Mie theory breaks down.

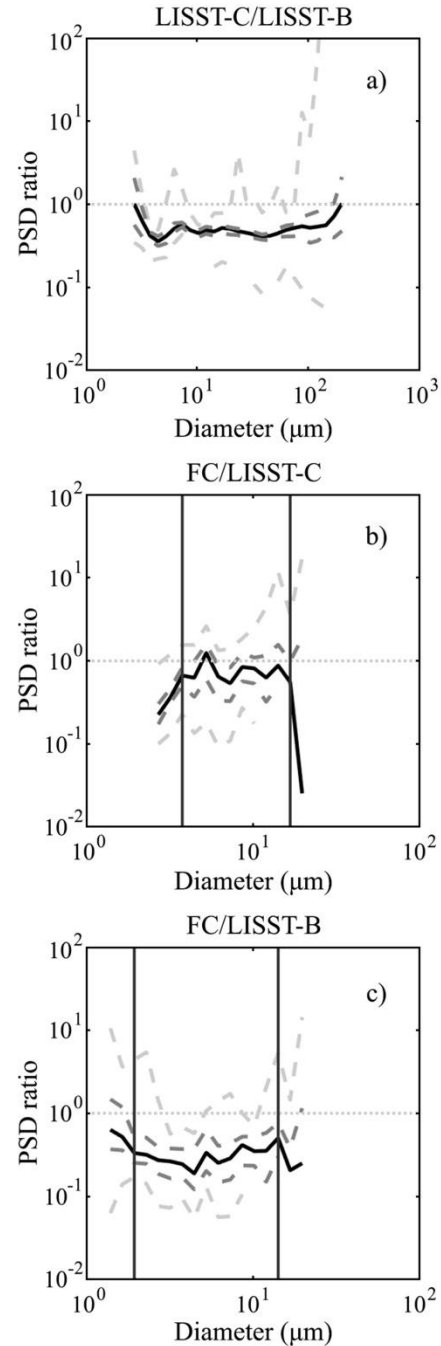


Fig. 9. Comparison between PSDs produced by the FC method and the two LISST-100x instruments used in this study, given as the median of all PSD/PSD ratios for the (a) LISST-C/LISST-B, (b) FC/LISST-C and (c) FC/LISST-B relationships (quartiles and max./min. given as dark and light grey dashed lines respectively). The vertical lines indicate the size range over which slopes for the LISST PSDs were calculated.

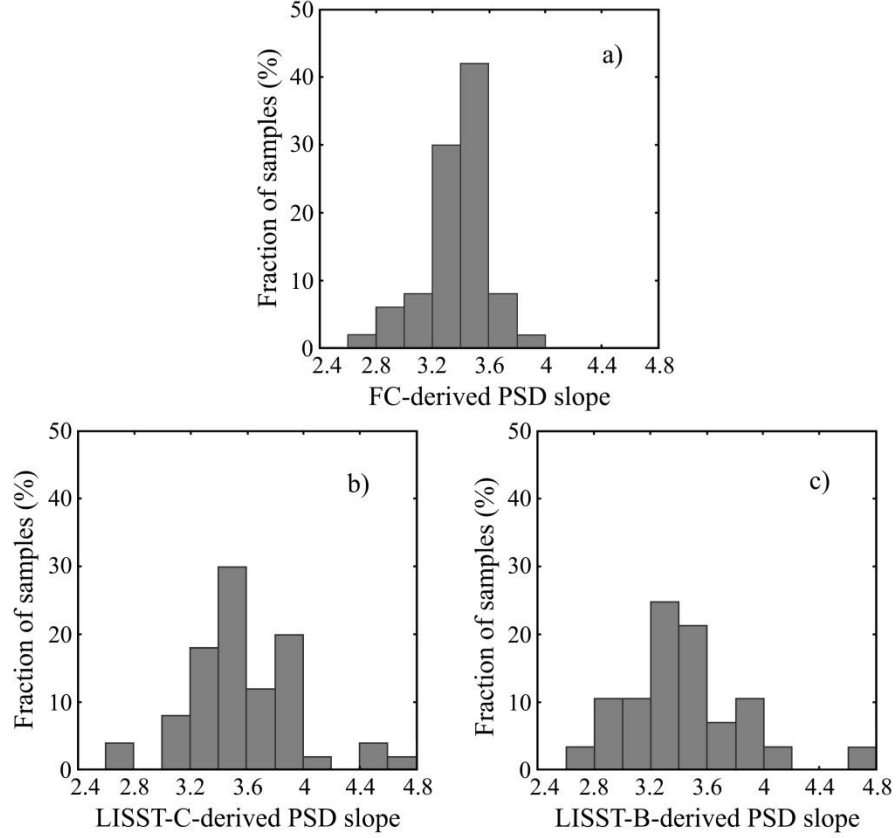


Fig. 10. Histograms of PSD slope for (a) FC PSDs, (b) LISST-C PSDs and (c) LISST-B PSDs. Note that LISST-B PSDs were available for only 28 out of the 50 samples of the UKCW dataset.

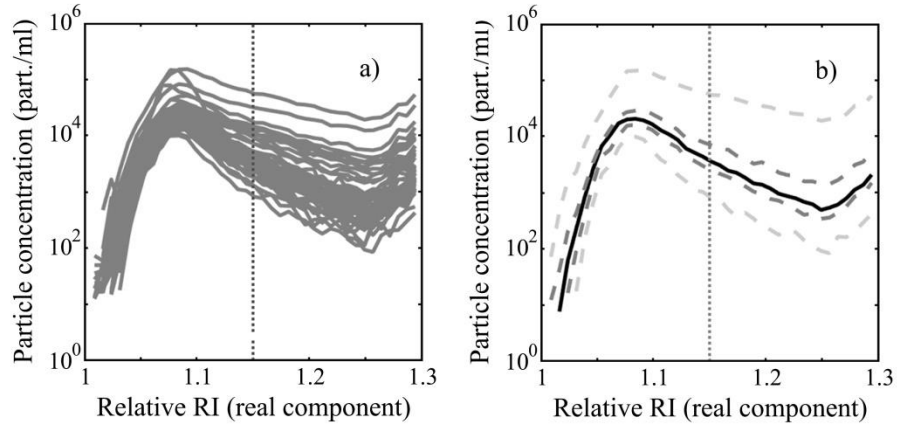


Fig. 11. (a) PRIDs retrieved by the FC method for the 50 samples of the UKCW dataset and (b) median of all 50 samples of the dataset (quartiles and max./min. given as dark and light grey dashed lines respectively). Real refractive index values above 1.15 (dotted line) are not precise, but still indicate high refractive indices.

At the number densities level, separating the particle population of a sample at the $n_r = 1.1$ threshold (taken as the midpoint between 1.05 and 1.15 as typical organic and inorganic n_r values) allowed for broad assessment of population dynamics (Fig. 12a). Samples collected in the Bristol Channel and the Irish Sea showed a prevalence of inorganic particles, while samples collected west of Orkney and in the North Sea were dominated by organics and samples from Loch Fyne and the Hebrides displayed a roughly balanced mixtures of organics and inorganics (Fig. 12b).

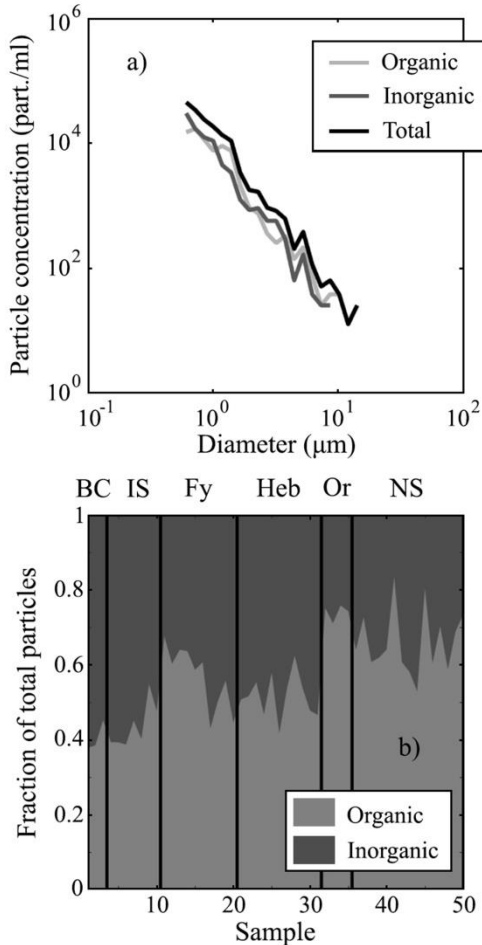


Fig. 12. (a) Total, organic ($n_r < 1.1$) and inorganic ($n_r \geq 1.1$) PSDs corresponding to the sample shown in Fig. 8a. (b) Organics vs. inorganics balance across the UKCW dataset. The vertical lines reflect the regional groupings of Fig. 6.

4. DISCUSSION

The FC method was found to be capable of identifying diameter and real refractive index within a range of diameters and n_r values in a variety of polystyrene standard bead samples and oils suspensions (Fig. 4-5). Diameters and real refractive indices were successfully identified in a range from ~ 0.5 to $\sim 10 \mu\text{m}$ for diameter and up to 1.15 for n_r . Due to the n_r isolines of the scattering grid bundling up together at high n_r values, real refractive indices couldn't be identified reliably for $n_r > \sim 1.15$, but still provided a generic high- n_r signifier as opposed to the more specific low n_r values.

When applied to natural water samples, the FC method produced PSDs which compared well with the features and slope of independent LISST-100x PSD measurements (Fig. 8-10). However, there was a lack

of consensus on particle concentration between the two LISSTs and the FC (Fig. 9). Independent calibration of the latter against particle concentration standards adds weight to the FC data, but further work is required to fully establish the reliability of concentration data for natural samples. Natural FC PRIDs were found to be largely homogeneous with a broad maximum between 1.05-1.15 and peak values at 1.07-1.09, consistent with a mixture of organic and inorganic particles (Fig. 11). The range of values accounts for most of the common components of marine suspended particulate matter [18,19].

A tail in the particle distribution with $n_r > 1.15$ is observed in all UKCW samples, with median concentrations 5 to 10 times lower than peak ones (Fig. 11). A concentration spike visible at ~ 1.3 in particular is found to be made up by particles outside the scattering model grid, which are assigned to the highest n_r considered, i.e. 1.3. These were found to account for less than 1% of the total particle population in all samples of the UKCW dataset, and interpreted as small, faceted, high- n_r particles incompatible with Mie theory assumptions. Due to their relatively small number the influence of particles outside the scattering model grid is negligible. Particles larger than $10 \mu\text{m}$ were also found to be exceedingly few in number in the UKCW samples, although this is likely to be an effect of the short acquisition times in the measurement protocol, which selected for smaller particles. Finding statistically significant numbers of large particles will necessitate a revision of the measurement protocol and at the very least longer acquisition times.

The PRIDs determined by the FC method offered the opportunity for a broad assessment of the population dynamics within the UKCW dataset, identifying geographical trends in the prevalence of organic and inorganic fractions of particles within the total PSDs (Fig. 12). In perspective, the PSD and PRID information retrieved by the FC method may also potentially be used to inform optical models such as Mie theory and biogeochemistry models of density and of organic carbon and chlorophyll content, estimating fractional contributions to bulk particulate IOPs, total suspended matter, total particulate organic carbon and chlorophyll.

5. CONCLUSION

The results found in this study demonstrated the potential of a flow cytometric method, which is unique in its capability of concurrently retrieving both size and refractive index of suspended marine particles on an individual basis in controlled samples with monotypic spherical particles (polystyrene beads, oil suspensions) and in natural water samples containing mixed populations. The resulting ability to partition the PSD by particle type not only offers the capability to resolve changes in particle population dynamics in natural water; the information that the FC method provides in terms of PSD and PRID determination, coupled with the inherent fluorescence detection capabilities of flow cytometry, represents a large and unexploited set of data which, once coupled with adequate models, may straddle physical, optical and biogeochemical particulate properties.

Moving forward, the PSDs and PRIDs retrieved by the FC method may therefore be used as inputs for Mie-based forward modelling, reconstructing IOP values and offering a chance of comparison with independent IOP measurements obtained by instruments such as spectrophotometers and backscattering meters, and for biogeochemistry modelling of total suspended matter and of carbon and chlorophyll concentrations. Agreement between measured and modelled values will further validate the FC method as a robust technique for diameter and n_r retrieval and also be an important step towards general optical closure. Furthermore, it will offer a chance to explore the biogeochemistry of particle populations, possibly tying together the optics, biology and geology of marine particles.

Funding Information. The HE442 cruise with RV Heincke was conducted under the grant number AWI-HE442. Agagiate received funding from the MASTS pooling initiative (The Marine Alliance for Science and Technology for Scotland). MASTS is funded by the Scottish Funding Council (grant reference HR09011) and contributing institutions.

Acknowledgment. The authors wish to thank the captain and the crew of RV Heincke for their support and help during the HE442 research cruise. MASTS funding support is gratefully acknowledged.

References

1. E. J. Davies, D. McKee, D. Bowers, G.W. Graham and W. A. M. Nimmo-Smith, "Optically significant particle sizes in seawater," *Applied Optics* 53, 1067 (2014).
2. S. G. Ackleson and R. W. Spinrad, "Size and refractive index of individual marine particulates: a flow cytometric approach," *Applied Optics* 27, 1270 (1988).
3. G. B. J. Dubelaar and R. R. Jonker, "Flow cytometry as a tool for the study of phytoplankton," *Scientia Marina* 64, 135-156 (2000).
4. C. M. Mikulski, S. L. Morton and G. J. Doucette, "Development and application of LSU rRNA probes for *Karenia brevis* in the Gulf of Mexico, USA," *Harmful Algae* 4, 49-60 (2005).
5. H. M. Sosik and R. J. Olson, "Automated taxonomic classification of phytoplankton sampled with imaging-in-flow cytometry," *Limnology and Oceanography: Methods* 5, 204-216 (2007).
6. R. J. Olson and H. M. Sosik, "A submersible imaging-in-flow instrument to analyze nano-and microplankton: Imaging FlowCytobot," *Limnology and Oceanography: Methods* 5, 195-203 (2007).
7. M. Thyssen, B. Beker, D. Ediger, D. Yilmaz, N. Garcia and M. Denis, "Phytoplankton distribution during two contrasted summers in a Mediterranean harbour: combining automated submersible flow cytometry with conventional techniques," *Environmental Monitoring and Assessment* 173, 1-16 (2010).
8. M. Thyssen, G. A. Tarran, M. Z. Zubkov, R. J. Holland, G. Gregori, P. H. Burkill and M. Denis, "The emergence of automated high-frequency flow cytometry: revealing temporal and spatial phytoplankton variability," *Journal of Plankton Research* 30, 333-343 (2007).
9. M. L. Brosnahan, S. Farzan, B. A. Keafer, H. M. Sosik, R. J. Olson and D. M. Anderson, "Complexities of bloom dynamics in the toxic dinoflagellate *Alexandrium fundyense* revealed through DNA measurements by imaging flow cytometry coupled with species-specific rRNA probes," *Deep Sea Research Part II: Topical Studies in Oceanography* 103, 185-198 (2014).
10. R. E. Green, H. M. Sosik and R. J. Olson, "Contributions of phytoplankton and other particles to inherent optical properties in New England continental shelf waters," *Limnology and Oceanography* 48, 2377-2391 (2003).
11. R. E. Green, H. M. Sosik, R. J. Olson and M. D. DuRand, "Flow cytometric determination of size and complex refractive index for marine particles: comparison with independent and bulk estimates," *Applied Optics* 42, 526 (2003).
12. L. Duforêt-Gaurier, W. Moutier, N. Guiselin, M. Thyssen, G. Dubelaar, X. Mériaux, L. Courcot, D. Dessailly and H. Loisel, "Determination of backscattering cross section of individual particles from cytometric measurements: a new methodology," *Optics Express* 23, 31510 (2015).
13. W. Moutier, L. Duforêt-Gaurier, M. Thyssen, H. Loisel, X. Mériaux, L. Courcot, D. Dessailly and S. Alvain, "Scattering of individual particles from cytometry: tests on phytoplankton cultures," *Optics Express* 24, 24188 (2016).
14. G. Mie, "Beiträge zur Optik trüber Medien, speziell kolloidaler Metallösungen," *Ann. Phys.* 25, 377-445 (1908).
15. H. C. van de Hulst, *Light Scattering By Small Particles* (Dover Publications, 2012).
16. C. F. Bohren and D. R. Huffman, *Absorption And Scattering Of Light By Small Particles* (Wiley, 2013).
17. W. H. Slade, "FASTMIE," <http://www.scattport.org/www.scattport.org/index.php/light-scattering-software/mie-type-codes/list/264-fastmie.html>. (2006)
18. E. Aas, "Refractive index of phytoplankton derived from its metabolite composition," *Journal of Plankton Research* 18, 2223-2249 (1996).
19. M. S. Twardowski, E. Boss, J. B. Macdonald, W. S. Pegau, A. H. Barnard and J. R. V. Zaneveld, "A model for estimating bulk refractive index from the optical backscattering ratio and the implications for understanding particle composition in case I and case II waters," *Journal of Geophysical Research: Oceans* 106, 14129-14142 (2001).
20. Y. C. Agrawal and H. C. Pottsmith, "Laser diffraction particle sizing in STRESS," *Continental Shelf Research* 14, 1101-1121 (1994).
21. P. Traykovski, R. Latter and J. D. Irish, "A laboratory evaluation of the laser in situ scattering and transmissometry instrument using natural sediments," *Marine Geology* 159, 355-367 (1999).
22. Y. C. Agrawal and H. C. Pottsmith, "Instruments for particle size and settling velocity observations in sediment transport," *Marine Geology* 168, 89-114 (2000).
23. Sequoia Scientific, Inc., *LISST-100X Particle Size Analyzer, User's Manual* (version 5.0) (2013)